FernUniversität in Hagen

LOTSE

Aufgabenteil zur Klausur zum

Modul 2.1 im BA-Studiengang „Politik- und Verwaltungswissenschaft“

und zum

Modul 3 im BA-Studiengang „Soziologie“

Termin: **6. September 2011, 14.00 - 18.00 Uhr**

Prüfer: **apl. Prof. Dr. H.-J. Mittag**

Abzugeben in einem ausgefüllten Klausurumschlag **ist** am Ende **nur der** maschinenauswertbare **Markierungsbogen**. Den Umschlag bitte nicht zukleben. Das Aufgabenheft, die Formelsammlung und das Konzeptpapier werden *nicht* eingesammelt.

Hinweise zur Bearbeitung der Klausur:

1. Bitte lesen Sie diese Hinweise vollständig und aufmerksam durch, bevor Sie mit der Bearbeitung beginnen. Legen Sie für die Identitätskontrolle Ihren Personalausweis und die Anmeldebestätigung neben die Klausurunterlagen. Während der Klausur sind das Rauchen und die Benutzung von Mobiltelefonen und anderen Geräten, die eine Verbindung zum Internet herstellen können, strikt untersagt. **Bitte schalten Sie Ihr Mobiltelefon aus**!

2. Die Klausur besteht aus 23 Aufgaben und zwar 15 Multiple-Choice-Aufgaben (Antwort-Auswahl-Verfahren) mit insgesamt 75 Punkten und 8 numerischen Aufgaben mit insgesamt 25 Punkten. Die Klausurdauer beträgt 240 Minuten.

3. Bitte kontrollieren Sie sofort, ob Sie ein vollständiges Klausurexemplar mit einem **Aufgabenteil auf weißem Papier**, eine **Formelsammlung mit Glossar und angehängtem Konzeptpapier auf gelbem Papier** sowie einen **LOTSE-Markierungsbogen** erhalten haben. Bevor Sie mit der Bearbeitung der Aufgaben beginnen, füllen Sie bitte den Identifikationsteil des Markierungsbogens aus. Tragen Sie dort Ihren Namen, Ihre Matrikelnummer, Ihre Anschrift sowie das Datum ein und unterschreiben Sie.

4. Für die Bewertung der Klausur sind ausschließlich Ihre Markierungen auf dem LOTSE-Markierungsbogen ausschlaggebend. Sie können nach Auswertung der Klausur online über das LOTSE-Korrektursystem einsehen, was Sie auf dem Markierungsbogen eingetragen haben und wie das System Ihre Eintragungen bewertet hat. Sie können so Ihre Antworten auch später mit der Musterlösung vergleichen.

5. Erfahrungen haben gezeigt, daß Sie spätestens 20 Minuten vor Abgabe der Klausur mit dem Markieren beginnen sollten. Kontrollieren Sie ganz am Schluss noch einmal Ihre Markierungen, bevor Sie den Markierungsbogen abgeben.

6. Bei jeder Aufgabe ist die maximal erreichbare Anzahl der Punkte angegeben. Insgesamt können Sie 100 Punkte erreichen. Bei Erreichen von 50 Punkten haben Sie die Klausur auf jeden Fall bestanden.

7. Sind die *numerischen Aufgaben* richtig beantwortet, erhalten Sie die volle Punktzahl, ansonsten werden i. d. R. 0 Punkte vergeben.

8. Bei den *Multiple-Choice-Aufgaben* sind **fünf Aussagen** vorgegeben, **von denen mindestens eine zutreffend ist**. Zutreffende Aussagen sind von Ihnen auf dem Markierungsbogen mit einem Strich, einem Kreuz oder einem Kreis zu kennzeichnen, falsche Aussagen sind nicht zu markieren. Wichtig ist, dass Ihre Markierungen nicht zu dünn sind und nicht in Nachbarfelder hineinreichen. Die Markierungen sind mit einem weichen Bleistift durchzuführen (empfohlen, weil von Ihnen noch änderbar) oder einem schwarzen Filzstift mittlerer Stärke. Falls Sie bei Verwendung eines Filzstifts noch Korrekturen vornehmen, müssen diese eindeutig und klar sein, damit wir sie anerkennen können. In der nachstehenden Grafik ist angedeutet, wie die Markierungen aussehen bzw. nicht aussehen sollten (die Eintragungen wurden hier willkürlich vorgenommen).

9. Die Bewertung der *Multiple-Choice-Aufgaben* erfolgt nach folgendem Prinzip: Sie erhalten

    - 1 Punkt, wenn Sie 3 der 5 vorgegebenen Antworten richtig haben,
    - 3 Punkte, wenn Sie 4 der 5 vorgegebenen Antworten richtig haben,
    - 5 Punkte, wenn Sie alle 5 Antworten richtig haben.

    Aufgaben, bei denen Sie weniger als 3 Antworten richtig haben, werden mit 0 Punkten bewertet. Aufgaben, bei denen Sie keine Markierung vornehmen, gelten als nicht bearbeitet und werden ebenfalls mit 0 Punkten bewertet. Das Verfahren berücksichtigt, dass bei geschlossenen Aufgabenformen schon durch bloßes Raten richtige Antworten erreicht werden können (siehe dazu http://www.fernuni-hagen.de/mks/lotse/gesamtbewertung.shtml).

10. Beispiel zur Bewertung der MC-Aufgaben: Sind die Aussagen A und B richtig sowie C, D und E falsch und es wurden A, B und C als richtig markiert, gibt es 3 Punkte, weil die Antworten zu A, B, D und E zutreffen.

11. Für Zwischenrechnungen können Sie das der Formelsammlung angehängte Konzeptpapier verwenden. Zwischenrechnungen gehen nicht in die Bewertung ein, weil nur der LOTSE-Bogen eingesammelt und verarbeitet wird.

12. Als Hilfsmittel ist neben der ausgeteilten **Formelsammlung** (mit Glossar) nur ein **Taschenrechner** zugelassen. Dieser darf nicht programmierbar sein und auch nicht über eine alphanumerische Tastatur verfügen. Ferner darf er keine Texte oder Formeln speichern und nicht drahtlos mit anderen Geräten kommunizieren können.

13. Täuschungen, Täuschungsversuche und andere Verstöße gegen die Prüfungsdisziplin können zum Ausschluss von der Klausur und zur Bewertung mit „nicht ausreichend (5,0)" führen.

14. Alle Klausurteilnehmer erhalten von der FernUniversität eine Benachrichtigung, auf dem die erreichte Punktzahl und die Note vermerkt sind. Die maschinelle Auswertung nimmt erfahrungsgemäß einen Zeitraum von 8 Wochen in Anspruch. Sehen Sie daher bis Ende Oktober von Nachfragen zum Klausurergebnis ab.

Viel Erfolg bei der Klausurbearbeitung!

Multiple-Choice-Aufgaben

Aufgabe 1 (Aussagenlogik) (5 Punkte)

In der nachstehenden Wahrheitstabelle stehen im Tabellenkopf der ersten beiden Spalten zwei Ausagen a und b, die wahr (w) oder falsch (f) sind. In den beiden ersten Spalten sind alle möglichen Fälle bezüglich des Wahrheitsgehalts der beiden Aussagen angegeben (beide Aussagen „w“, nur eine Aussage „w“, beide Aussagen „f“).

a	b	P1: $a \wedge b$	P2: $\neg b$	K: $(a \wedge b) \vee (\neg b)$
w	w			
w	f			
f	w			
f	f			

Bei den anderen Spalten ist nur der Tabellenkopf ausgefüllt. Hier sind zwei Prämissen $P1$ und $P2$ angegeben, die sich aus a und b ableiten. Im Kopf der letzten Spalte steht ein aus den Prämissen abgeleiteter logischer Schluss K, der allerdings noch auf seinen Wahrheitsgehalt zu überprüfen ist. Das Zeichen $\neg$ bedeutet die Negation einer Aussage, $\wedge$ (Konjunktion) beinhaltet ein logisches „und“ (zwei Aussagen gelten gleichzeitig), während das Zeichen $\vee$ (Disjunktion) ein logisches „oder“ darstellt (von zwei Aussagen gilt mindestens eine).

Welche der folgenden Aussagen sind richtig ? (x aus 5)

A) Wenn die Aussagen a und b beide wahr sind, sind auch die Prämissen $P1$ und $P2$ erfüllt, also beide wahr.

B) Wenn die Aussagen a und b beide falsch sind, sind sowohl $P1$ als auch $P2$ nicht erfüllt.

C) Wenn die Aussage a wahr und b falsch ist, ist $P2$ erfüllt, nicht aber $P1$.

D) Wenn die Aussage a falsch und b wahr ist, ist $P2$ erfüllt, nicht aber $P1$.

E) Wenn eine der beiden Prämissen $P1$ und $P2$ erfüllt ist, ist auch K wahr, d. h. die Konklusion ist dann korrekt.

Hinweis: Ergänzen Sie am besten zunächst die 12 fehlenden Werte der obigen Tabelle, d. h., setzen sie jeweils w oder f ein. Die ausgefüllte Tabelle geht zwar nicht in die maschinelle Bewertung ein, erleichtert es Ihnen aber, die zutreffenden der fünf vorgegebenen Antwortalternativen zu finden.

Aufgabe 2 (Merkmalsklassifikationen) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Der Begriff „metrische Skala“ ist als Oberbegriff für „Intervallskala“ und „Verhältnisskala“ (letztere einschließlich des Sonderfalls „Absolutskala“) zu verstehen. (x aus 5)

A) Bei einem nominalskalierten Merkmal lassen sich die Merkmalsausprägungen nicht in eine Rangfolge bringen.

B) Operationen, die für metrisch skalierte Daten zulässig sind, sind ebenso für ordinalskalierte Daten zulässig.

C) Das Merkmal „Bei einer Landtagswahl gewählte Partei“ ist ein nominalskaliertes Merkmal.

D) Metrisch skalierte Merkmale sind stets quantitativ.

E) Die Anzahl der im zweiten Quartal 2011 pro Tag in Hamburg neu gemeldeten EHEC-Fälle (EHEC = bakterielle Darmerkrankung) ist ein stetiges Merkmal.

Aufgabe 3 (Informationsgewinnung, Datenerhebung) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Bei den Aussagen B, C, D und E geht es jeweils um die Beurteilung der Richtigkeit des letzten Satzes. (x aus 5)

A) Bei einer Befragung (per Fragebogen oder Interview) dienen Filterfragen dem Zweck, Untergruppen von Befragten zu bilden, die dann jeweils spezielle Fragen zu beantworten haben.

B) Bei der Gestaltung von Fragebögen kann es vorkommen, dass eine Frage die Antwort auf eine nachfolgende Frage beinflusst. Eine solche Ausstrahlungswirkung nennt man Halo-Effekt.

C) Um theoretische Konstrukte (z. B. „Schulischer Erfolg“ oder „Lebenszufriedenheit“) messen zu können, muss man sie mit beobachtbaren Konstrukten verknüpfen, d. h. es gilt Handlungsanweisungen für die Gewinnung von Daten zu spezifizieren. Diesen Schritt bezeichnet man als Operationalisierung.

D) Anhand des Random-Route-Verfahrens lassen sich Zufallsstichproben von Personen aus einer Zielpopulation gewinnen. Die Anwendung des Verfahrens setzt das Vorhandensein einer Namens- oder Adressdatei voraus.

E) Bei Experimenten mit Personen, z. B. in der Psychologie, werden i. a. zwei Gruppen von Teilnehmern gebildet – eine Versuchsgruppe und eine Kontrollgruppe. Anschließend werden in einer Gruppe – der Versuchsgruppe – Einflussgrößen planmäßig verändert.

Aufgabe 4 (Messen / Stichprobenverfahren) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Bei Aufgabenteil D geht es um die Beurteilung der Richtigkeit des letzten Satzes. (x aus 5)

A) Die Reliabilität charakterisiert, inwieweit ein Messinstrument bei wiederholter Messung die gleichen Messwerte liefert.

B) Aus der Reliabilität einer Messung folgt stets auch deren Validität.

C) Die Klumpenauswahl ist eine zufallsgesteuerte Auswahlprozedur, bei der sich die Zufallsauswahl auf Teilmengen einer Grundgesamtheit bezieht, nicht auf die Untersuchungseinheiten selbst.

D) Bei einer geschichteten Stichprobenauswahl wird eine Grundgesamtheit in Teilmengen zerlegt (sog. Schichten), denen dann jeweils Zufallstichproben entnommen werden. Dabei wird jeder Schicht stets ein fester Prozentsatz von Stichprobenelementen zufällig entnommen.

E) Die Quotenauswahl ist ein zweistufiges Stichprobenverfahren, bei dem auf der zweiten Verfahrensstufe keine zufällige, sondern eine systematische Auswahl der Stichprobenelemente erfolgt.

Aufgabe 5 (Zusammenhangsmessung) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Die Aussagen A - D beziehen sich auf empirische Zusammenhangsmaße, sind also aus Beobachtungsdaten errechenbar. Aussage E bezieht sich hingegen auf ein theoretisches Zusammenhangsmaß, also auf ein Zusammenhangsmaß für Zufallsvariablen. (x aus 5)

A) Der Korrelationskoeffizient r nach Bravais-Pearson misst die Stärke eines linearen Zusammenhangs zwischen zwei Merkmalen X und Y.

B) Wenn $r = 1$ ist, bedeutet dies, dass die Datenpaare $(x_1, y_1), ..., (x_n, y_n)$ alle auf einer steigenden oder fallenden Geraden liegen.

C) Im Falle $r = 0$ ist noch nicht ausgeschlossen, dass zwischen den Merkmalen X und Y ein nicht-linearer Zusammenhang besteht.

D) Der Rangkorrelationskoeffizient r_{SP} lässt sich – unter Informationsverlust – auch auf metrisch skalierte Merkmale anwenden.

E) Die Kovarianz zweier Zufallsvariablen X und Y ist ein Zusammenhangsmaß, das nur Werte zwischen -1 und $+1$ annehmen kann.

Aufgabe 6 (Univariate Häufigkeitsverteilungen / Kenngrößen) (5 Punkte)

Gegeben sei der folgende Datensatz für ein stetiges Merkmal X:
$4,8 \quad 6,4 \quad 4,2 \quad 4,6 \quad 4,8 \quad 3,9 \quad 4,2 \quad 7,6 \quad 6,5.$

Welche der folgenden Aussagen, die alle von diesem Datensatz ausgehen, sind richtig? Beachten Sie, dass eine aus mehreren Teilaussagen bestehende Aussage nur dann als richtig zu bewerten ist, wenn jede Teilaussage zutrifft. (x aus 5)

A) Der obige Datensatz hat einen eindeutig bestimmten Modalwert.

B) Der Median $\widetilde{x}$ des obigen Datensatzes ist kleiner als dessen Mittelwert $\overline{x}$.

C) Wenn man bei obigem Datensatz den letzten Wert $(6,5)$ der Urliste um $1,2$ erhöht, hat dies zur Folge, dass sowohl der Mittelwert $\overline{x}$ als auch der Median $\widetilde{x}$ des Datensatzes größer werden.

D) Mit der in Aufgabenteil C spezifizierten Veränderung des letzten Wertes $(6,5)$ der Urliste verändert sich die Spannweite des Datensatzes.

E) Wenn man bei dem eingangs aufgeführten Datensatz den letzten Wert $(6,5)$ streicht, werden sowohl der Mittelwert $\overline{x}$ als auch der Median $\widetilde{x}$ des Datensatzes kleiner.

Aufgabe 7 (Konzentrationsmessung) (5 Punkte)

Bei einer Aktiengesellschaft verteilt sich der Aktienbesitz im Gesamtwert von 20 Millionen Euro auf nur 4 Aktionäre. Aktionär 1 besitzt mit 10 % des Aktienpakets den kleinsten, Aktionär 2 mit 40 % den größten Anteil. Die beiden Aktionäre 3 und 4 sind mit je 25 % beteiligt. Der Aktienbesitz ist also innerhalb der kleinen Aktonärsgruppe nicht gleichmäßig verteilt.

Welche der folgenden Aussagen sind richtig ? (x aus 5)

A) Wenn man auf der Basis des obigen Datensatzes die Lorenzkurve zeichnet – also einen Polygonzug, der den Nullpunkt mit den Punkten $(0,25; v_1)$, $(0,5; v_2)$, $(0,75; v_3)$ und $(1; 1)$ verbindet – so nimmt diese Kurve an der Stelle $0,5$ einen Wert an, der zwischen $0,30$ und $0,34$ liegt.

B) Der Wert, den man für den Ordinatenwert v_3 errechnet, gibt an, welcher Anteil des gesamten Aktienpakets auf diejenigen 3 Aktionäre entfällt, die die kleinsten Anteile am gesamten Aktienpaket halten.

C) Allgemein gilt, dass der unnormierte Gini-Koeffizient für einen Datensatz des Umfangs $n = 4$ den Wert $0,75$ nicht überschreiten kann.

D) Der normierte Gini-Koeffizient hat bei obigem Datensatz einen Wert G^*, der zwischen $0,29$ und $0,31$ liegt.

E) Wenn Aktionär 1 seinen Anteil an Aktionär 2 verkaufte, Aktionär 1 also aus der kleinen Aktionärsgruppe ausschiede, würde der Wert des dann resultierenden normierten Gini-Koeffizienten G^* oberhalb von $0,31$ liegen.

Aufgabe 8 (Randverteilungen, bedingte Wahrscheinlichkeiten) (5 Punkte)

An der FernUniversität gab es im Sommersemester 2011 für einen noch relativ neuen BA-Studiengang insgesamt 128 Neueinschreibungen (Vollzeit- oder Teilzeitstudium). Es entschieden sich 100 Studierende für ein Teilzeitstudium. Ferner waren 48 der neu eingeschriebenen Studierenden männlich, wobei 12 Männer ein Vollzeitstudium wählten.

Welche der folgenden Aussagen, die sich auf die Einschreibhäufigkeiten für den genannten Studiengang beziehen, sind richtig? (x aus 5)

A) Von den weiblichen Neueinschreibern im Sommersemester hatten sich 20 % für ein Vollzeitstudium entschieden.

B) Wählt man aus den neu eingeschriebenen 128 Studierenden eine Person zufällig aus, so liegt die Wahrscheinlichkeit, dass diese im Vollzeitmodus studiert, über $0,22$.

C) Wählt man aus der Gruppe aller 128 Neueinschreiber des Sommersemesters eine Person zufällig aus, so liegt die Wahrscheinlichkeit, dass ein im Teilzeitmodus studierender Mann ausgewählt wird, über $0,27$.

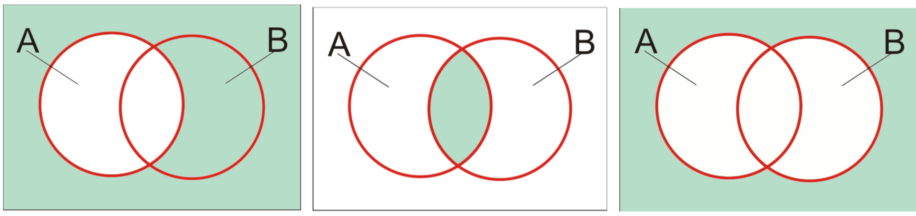
D) Die Wahrscheinlichkeit, dass bei Zufallsauswahl einer Frau die Wahl auf eine Frau fällt, die im Teilzeitmodus studiert, liegt zwischen $0,79$ und $0,81$.

E) Der Anteil der Frauen lag bei den Vollzeitstudierenden höher als bei den Teilzeitstudierenden.

Aufgabe 9 (Venn-Diagramme) (5 Punkte)

Zur Veranschaulichung von Ereignissen oder von Mengen lassen sich Venn-Diagramme heranziehen. Diese bestehen aus einem Rechteck, in dem die Ereignisse als Kreise oder Ellipsen dargestellt sind. Das Rechteck repräsentiert eine Grundgesamtheit, von der die eingezeichneten Mengen Teilmengen sind. Es bezeichnen $\overline{A}$ und $\overline{B}$ die Komplementärmengen von A und B, $A \cap B$ deren Schnittmenge und $A \cup B$ die Vereinigungsmenge von A und B. Zwei Mengen, deren Darstellungen in Venn-Diagrammen sich nicht überschneiden, werden als disjunkt bezeichnet.

Nachstehend sind drei Venn-Diagramme abgebildet, die sich auf die Verknüpfung zweier durch Kreise dargestellten Ereignisse oder Mengen A und B beziehen:



Welche der folgenden Aussagen, die auf die obigen Diagramme Bezug nehmen, sind richtig? (x aus 5))

A) Das erste Venn-Diagramm (von links gezählt, also in der üblichen Leserichtung) veranschaulicht anhand der dunkler gefärbten Fläche die Komplementärmenge $\overline{A}$ von A.

B) Im zweiten Venn-Diagramm ist durch die dunkler gefärbte Fläche die Schnittmenge von A und B dargestellt, also $A \cap B$.

C) Im dritten Venn-Diagramm ist durch die dunkler gefärbte Fläche die Vereinigungsmenge der Komplementärmengen von A und B dargestellt, also $\overline{A} \cup \overline{B}$.

D) Die beiden Komplementärmengen $\overline{A}$ und $\overline{B}$ sind disjunkt.

E) Die Vereinigung der beiden Schnittmengen $A \cap B$ und $A \cap \overline{B}$ liefert A.

Aufgabe 10 (diskrete Zufallsvariablen) (5 Punkte)

Es sei eine diskrete Zufallsvariable X mit 10 Ausprägungen gegeben, nämlich den Ausprägungen $x_1 = 0, x_2 = 1, \ldots, x_{10} = 9$. Alle 10 Ausprägungen besitzen dieselbe Eintrittswahrscheinlichkeit.

Welche der folgenden Aussagen sind richtig? (x aus 5)

A) Die Wahrscheinlichkeitsfunktion $f(x)$ der diskreten Zufallsvariablen X nimmt für $x = 1$ den Wert $0,1$ an.

B) Die Verteilungsfunktion $F(x)$ von X nimmt für $x = 0$ den Wert 0 an.

C) Die Verteilungsfunktion $F(x)$ von X nimmt für $x = 1,5$ den Wert $0,2$ an.

D) Für den Erwartungswert $\mu = E(X)$ der Zufallsvariablen X gilt $\mu = 5$.

E) Die Varianz $V(Y)$ der durch $Y = 2 \cdot X$ definierten Zufallsvariablen Y ist doppelt so groß wie die Varianz von X.

Aufgabe 11 (Kombinatorik / diskrete Verteilungen) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Bei Aufgabenteil C geht es um die Beurteilung der Richtigkeit des letzten Satzes. (x aus 5)

A) Wenn man eine faire Münze, also eine Münze mit gleichen Eintrittswahrscheinlichkeiten für „Kopf" und „Zahl", 8-mal wirft und die Anzahl X der Ausgänge mit „Zahl" feststellt, liegt die Wahrscheinlichkeit dafür, *mindestens* viermal „Zahl" zu erhalten, zwischen $0,60$ und $0,62$.

B) Die Wahrscheinlichkeit dafür, bei dem 8-maligem Münzwurf aus Aufgabenteil A *genau* viermal „Zahl" zu erhalten, ist größer als $0,26$.

C) Mit einem fairen Würfel, also einem Würfel mit gleichen Eintrittswahrscheinlichkeiten für alle Augenzahlen, werde 10-mal in Folge gewürfelt. Bezeichne X die Anzahl der Ausgänge mit einer Augenzahl, die größer als 4 ist. Der Erwartungswert von X liegt zwischen $3,2$ und $3,4$.

D) Beim Würfeln mit zwei fairen Würfeln, liegt die Wahrscheinlichkeit dafür, eine Augensumme zu erhalten, die nicht kleiner als 10 ist, zwischen $0,30$ und $0,32$.

E) Der Erwartungswert für die Augensumme beim Würfeln mit zwei fairen Würfeln beträgt 7.

Aufgabe 12 (Punkt- und Intervallschätzungen) (5 Punkte)

Gegeben seien n Stichprobenwerte, die als Ausprägungen x_1, x_2, ... ,x_n unabhängiger Zufallsvariablen X_1, X_2, ... ,X_n interpretiert werden. Die Zufallsvariablen X_1, X_2, ... ,X_n seien alle als normalverteilt spezifiziert mit gleichem Erwartungswert μ und gleicher Varianz σ^2. Sowohl μ als auch σ^2 sollen geschätzt werden.

Zur Schätzung der beiden genannten Verteilungsparameter können z. B. der Stichprobenmittelwert $\overline{X}$ und die Stichprobenvarianz S^2 herangezogen werden. Beide genannten Stichprobenfunktionen leiten sich aus den Zufallsvariablen X_1, X_2, ... ,X_n ab, wobei S^2 definiert sei als Summe der quadrierten Mittelwertabweichungen $(X_i - \overline{X})^2$, dividiert durch n. Mit X_1, X_2, ... ,X_n sind auch $\overline{X}$ und S^2 Zufallsgrößen.

Welche der folgenden Aussagen sind richtig? (x aus 5)

A) Die Stichprobenfunktion $\overline{X}$ liefert eine unverzerrte Schätzung für den Erwartungswert μ.

B) Die Varianz der Zufallsvariablen X_1, X_2, ... ,X_n und die Varianz der Stichprobenfunktion $\overline{X}$ stimmen überein.

C) Die Stichprobenfunktion S^2 liefert eine verzerrte Schätzung für die Varianz σ^2.

D) Wenn man den Erwartungswert μ der zugrunde gelegten Normalverteilung nicht anhand des Stichprobenmittelwerts schätzt (Punktschätzung), sondern anhand eines Konfidenzintervalls (Intervallschätzung), so sind die Grenzen des Konfidenzintervalls zufallsabhängig.

E) Ein Konfidenzintervall für den Erwartungswert μ muss den zu schätzenden Parameter μ nicht notwendigerweise enthalten.

Aufgabe 13 (Stetige Verteilungen) (5 Punkte)

Welche der folgenden Aussagen sind richtig? Beachten Sie, dass eine aus mehreren Teilen bestehende Aussage nur dann als richtig gilt, wenn sie in allen Teilen zutrifft. Bei den Aussagen A, B und E geht es um die Beurteilung des Wahrheitsgehalts des jeweils letzten Satzes. (x aus 5)

A) Es sei X eine stetige Zufallsvariable mit der Verteilungsfunktion $F(x)$ und es bezeichne $x_{0,01}$ das $0,01$-Quantil der Verteilung. Die Wahrscheinlichkeit dafür, dass X eine Ausprägung x mit $x > x_{0,01}$ annimmt, beträgt $0,99$.

B) Es sei X eine stetige Zufallsvariable mit der Dichtefunktion $f(x)$ und der Verteilungsfunktion $F(x)$. Der Inhalt der Fläche unterhalb der Dichtefunktion bis zum Punkt $x = 1$ stimmt mit dem Wert $F(1)$ der Verteilungsfunktion an der Stelle $x = 1$ überein.

C) Die t-Verteilung ist eine bezüglich des Nullpunkts symmetrische Verteilung, deren Varianz mit zunehmender Anzahl der Freiheitsgrade zunimmt.

D) Sind X_1 und X_2 zwei stetige Zufallsvariablen mit Varianz $V(X_1)$ resp. $V(X_2)$ und bezeichnet $Y = X_1 + X_2$ ihre Summe (also eine aus X_1 und X_2 gebildetete neue Zufallsvariable), so gilt stets $V(Y) = V(X_1) + V(X_2)$.

E) Bei einer einfaktoriellen Varianzanalyse wird die Testentscheidung davon abhängig gemacht, ob die verwendete Testgröße das $0,95$-Quantil einer F-Verteilung mit $m = 4$ und $n = 10$ Freiheitsgraden überschreitet. Das für die Entscheidung benötigte Quantil hat hier den Wert $5,96$.

Aufgabe 14 (Korrelationsmessung, lineares Regressionsmodell) (5 Punkte)

In der nachstehende Tabelle sind für zwei Merkmale X und Y Beobachtungsdaten $(x_i; y_i)$ wiedergegeben $(i = 1, 2, ..., 5)$.

i	x_i	y_i
1	2,7	2,5
2	3,1	4,5
3	2,0	1,5
4	3,8	4,5
5	3,4	3,0

Aus diesen Daten errechnet man $\overline{x} = 3,0$ sowie $\overline{y} = 3,2$. Für die empirischen Varianzen s_x^2 und s_y^2 erhält man die Werte $s_x^2 = 0,38$ resp. $s_y^2 = 1,36$ und für die empirische Kovarianz $s_{xy} = 0,60$. Diese fünf Werte können hier ungeprüft übernommen werden.

Welche der folgenden Aussagen sind richtig? Bei den Aussagen B, D und E geht es um die Beurteilung des Wahrheitsgehalts des jeweils letzten Satzes. (x aus 5)

A) Für den obigen Datensatz errechnet man für den Korrelationskoeffizienten r nach Bravais-Pearson einen Wert zwischen $0,80$ und $0,82$.

B) Wenn man das lineare Regressionsmodell $y_i = \alpha + \beta x_i + u_i$ für obigen Datensatz $(x_i; y_i)$ heranzieht $(i = 1, 2, ..., 5)$, kann man die Regressionskoeffizienten nach der Methode der kleinsten Quadrate schätzen. Erhielte man dabei für β z. B. einen Schätzwert $\widehat{\beta} = 1,9$ (fiktiver Wert), beinhaltete dieses Ergebnis, dass bei einer Erhöhung von X um eine Einheit mit einen Anstieg des Wertes für das Merkmal Y um $1,9$ Einheiten zu rechnen wäre.

C) Wenn man für den oben aufgeführten Datensatz tatsächlich die Kleinst-Quadrat-Schätzung $\widehat{\beta}$ berechnet, resultiert ein Wert, der zwischen $1,55$ und $1,65$ liegt.

D) Zur Beurteilung der Anpassungsgüte der nach der Kleinst-Quadrat-Methode ermittelten Regressionsgeraden verwendet man das Bestimmtheitsmaß R^2. Im Falle $R^2 = 0$ kann gefolgert werden, dass zwischen X und Y weder ein linearer noch ein nicht-linearer Zusammenhang besteht.

E) Es seien wieder fünf Datenpaare $(x_1; y_1), ..., (x_5; y_5)$ gegeben. Dabei stamme $(x_1; y_1) = (2,7; 2,5)$ aus der vorstehenden Tabelle, während $(x_2; y_2), ..., (x_5; y_5)$ andere, hier nicht wiedergegebene Datenpaare seien. Auf der Basis dieser fünf Datenpaare seien für die Koeffizienten α und β des linearen Regressionsmodells nach der Methode der kleinsten Quadrate die Schätzungen $\widehat{\beta} = 0,80$ und $\widehat{\alpha} = 0,15$ bestimmt worden. Für das Residuum $\widehat{u}_1 = y_1 - \widehat{y}_1$ errechnet sich dann ein Wert, der kleiner als $0,18$ ist.

Aufgabe 15 (Gauß-Test, Fehler beim Testen, Gütefunktion) (5 Punkte)

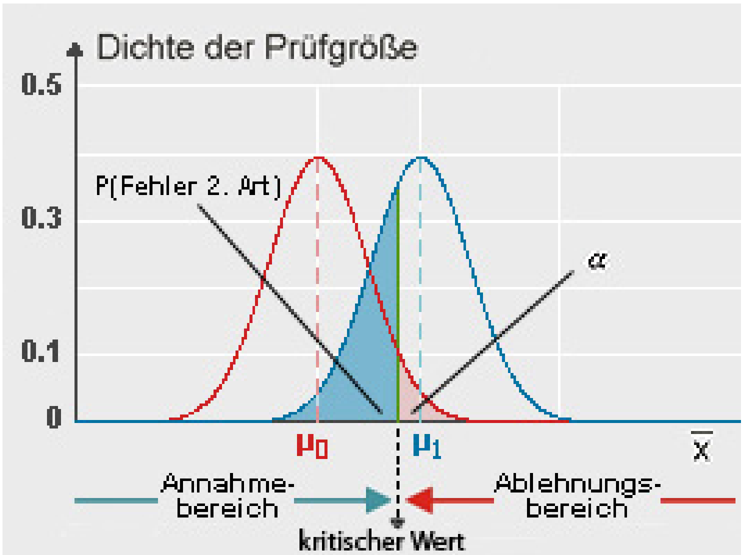
Es seien n unabhängige Beobachtungen für ein normalverteiltes Merkmal gegeben (Normalverteilung mit unbekanntem Erwartungswert μ und bekannter Varianz σ^2). Getestet werden soll

$$H_0: \ \mu \leq \mu_0 \qquad \text{gegen} \qquad H_1: \ \mu > \mu_0$$

und zwar zum Signifikanzniveau $\alpha = 0,05$ (Gauß-Test). Als Prüfstatistik kann der Stichprobenmittelwert $\overline{X}$ verwendet werden, der ebenfalls normalverteilt ist (mit gleichem Erwartungswert und kleinerer Varianz $\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}$), oder aber die nach Standardisierung resultierende Prüfgröße

$$Z = \frac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} = \frac{\overline{X} - \mu_0}{\sigma} \cdot \sqrt{n}.$$

Verwendet man die Prüfstatistik $\overline{X}$, so lässt sich deren Verteilung im Falle $\mu = \mu_0$, also wenn H_0 gerade noch zutrifft, durch eine Dichtefunktion charakterisieren, die in $\mu = \mu_0$ ihr Zentrum hat (linke Dichtekurve in der nachstehenden Grafik). Die Nullhypothese wird verworfen, wenn die Prüfgröße einen kritischen Wert überschreitet. Dieser Wert wird so bestimmt, dass er im Falle $\mu = \mu_0$ nur mit der vorgegebenen Wahrscheinlichkeit α überschritten wird.



Die obige Grafik zeigt noch eine zweite Dichtekurve für $\overline{X}$, die sich auf den Fall bezieht, dass H_1 zutrifft, der Erwartungswert μ also einen Wert $\mu = \mu_1$ besitzt, der größer als μ_0 ist. Unterhalb dieser zweiten Dichtekurve ist ebenfalls ein Flächenanteil markiert. Dieser veranschaulicht für den in der Grafik gewählten Wert μ_1 die Wahrscheinlichkeit des Eintritts eines Fehlers 2. Art.

Welche der folgenden Aussagen sind richtig? (x aus 5)

A) Die Wahrscheinlichkeit dafür, die Nullhypothese zu verwerfen, wenn $\mu < \mu_0$ gilt, ist kleiner als 0,05.

B) Ein Fehler 2. Art kann nur eintreten, wenn H_1 gilt, also nur für Werte μ mit $\mu > \mu_0$.

C) Die Wahrscheinlichkeit dafür, die Nullhypothese zu verwerfen, wenn $\mu = \mu_1$ gilt (s. Grafik), ist identisch mit dem Wert der Gütefunktion des Tests an der Stelle $\mu = \mu_1$.

D) Die Wahrscheinlichkeit für den Eintritt eines Fehlers 2. Art wird bei dem eingangs beschriebenen Test immer größer, je stärker der in der Grafik eingezeichnete Wert $\mu = \mu_1$ den Wert μ_0 überschreitet.

E) Wenn man den eingangs beschriebenen Test nicht mit der Prüfstatstistik $\overline{X}$, sondern unter Verwendung der standardisierten Prüfvariablen Z durchführt, ist der kritische Wert durch das $0,95$-Quantil der Standardnormalverteilung gegeben.

Numerische Aufgaben

Aufgabe 41 (Rangkorrelationskoeffizient) (3 Punkte)

Zwei Banken beurteilen unabhängig voneinander Anträge auf Gewährung eines Kredits, die von vier Existenzgründern eingereicht wurden. Basis für die Beurteilung des mit der Vergabe eines Kredits verbundenen Risikos sind jeweils die hauseigenen Bewertungsrichtlinien und die vorgelegten Businesspläne der Antragsteller.

Bei beiden Banken wird die Risikobewertung anhand einer 10-stufigen Ratingskala vorgenommen, wobei die Punktzahl 10 die beste Bewertung repräsentiert. Die Ergebnisse der Bewertungen sind nachstehend ausgewiesen.

Kreditantrag i	Bank A Bewertung x_i	Bank B Bewertung y_i
1	4	5
2	7	9
3	9	8
4	8	6

Untersuchen Sie anhand des Rangkorrelationskoeffizienten r_{SP} von Spearman, ob zwischen den Bewertungen der beiden Banken ein Zusammenhang besteht. Tragen Sie Ihr Ergebnis auf *zwei* Stellen nach dem Dezimalkomma genau rechtsbündig in das Antwortfeld ein. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch) $r_{SP} =$ | | | | | | |

Aufgabe 42 (Kombinatorik) (4 Punkte)

Aus einer Gruppe von 5 Personen, die aus 2 Frauen und 3 Männern besteht, werden im Rahmen eines Gewinnspiels zwei Gewinner ermittelt. Dazu wird jeder Person eine der Zahlen 1, 2, ..., 5 zugeordnet, die jeweilige Zahl auf einem Zettel notiert und die Zettel in identischen Briefumschlägen abgelegt. Nach Durchmischen der Umschläge werden nacheinander und ohne Zurücklegen zwei Umschläge zufällig ausgewählt. Die in den gezogenen Umschlägen enthaltenen Zahlen definieren dann die Gewinner.

Wie groß ist die Wahrscheinlichkeit dafür, dass keine der beiden Frauen einen Gewinn erhält? Tragen Sie Ihr Ergebnis, also einen Wert aus dem Intervall $[0; 1]$, auf *drei* Stellen nach dem Dezimalkomma genau rechtsbündig in das Antwortfeld ein. Das **Dezimalkomma** belegt ein **eigenes Feld**. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)

Aufgabe 43 (Gewinnwahrscheinlichkeiten bei Ziehung von Losen) (3 Punkte)

Aus einer Lostrommel mit 100 Losen, von denen 5 mit einem Gewinn verbunden sind, werden nacheinander 10 Lose *mit* Zurücklegen gezogen. Berechnen Sie die Wahrscheinlichkeit dafür, nach den 10 Ziehungen *höchstens einen* Gewinn gezogen zu haben.

Tragen Sie Ihr Ergebnis rechtsbündig und auf *vier* Nachkommastellen genau in das Antwortfeld ein, also z. B. $0,7748$ (nicht aber $77,48$ %). Das **Dezimalkomma** belegt ein **eigenes Feld**. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch)

Aufgabe 44 (Vorstandswahl mit Möglichkeit der Stimmenhäufung) (3 Punkte)

Bei einem größeren Konzern sollen nach Ausscheiden von zwei Vorstandsmitgliedern zwei Personen in den Vorstand nachrücken. Zur Auswahl stehen vier Kandidaten, nämlich Frau B, Herr D, Herr F und Herr S. Die Mitglieder des Auswahlgremiums setzen bei einer geheimen Wahl auf dem Wahlzettel 2 Kreuze, wobei zwei verschiedene Bewerber je einmal oder ein Bewerber zweimal angekreuzt werden kann (Möglichkeit der Stimmenhäufung).

Wieviele Möglichkeiten der Auswahl von 2 Bewerbern gibt es? Beachten Sie: Zwischen den Wahlausgängen (Kandidat x; Kandidat y) und (Kandidat y; Kandidat x) muss hier nicht unterschieden werden – sie führen ja zur selben Zusammensetzung des neuen Vorstands.

Tragen Sie Ihr Ergebnis, also eine positive ganze Zahl, rechtsbündig in das Antwortfeld ein. Übertragen Sie Ihr Ergebnis rechtzeitig vor Ende der Klausur auf den Markierungsbogen.

(numerisch) | | | | | | | |

Aufgabe 45 (Schwellenwertbestimmung bei einer Normalverteilung) (3 Punkte)

Es sei eine Zufallsvariable X betrachtet, die normalverteilt ist mit Erwartungswert $\mu = 1$ und Varianz $\sigma^2 = 0,25$. Bestimmen Sie einen Wert a, für den $P(X > a) = 0,8$ gilt, also einen Wert a, der mit Wahrscheinlichkeit $0,8$ überschritten wird.

Geben Sie das Ergebnis auf *vier* Stellen nach dem Dezimalkomma genau an. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**. Falls Sie also z. B. $1,935$ errechnen, tragen Sie in die letzten sechs Felder $1,9350$ ein. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch) $a =$ | | | | | | | |

Aufgabe 46 (Wahrscheinlichkeiten bei Normalverteilung) (3 Punkte)

Es sei angenommen, dass ein Merkmal X normalverteilt sei und zwar mit Erwartungswert $\mu = 500$ und einer Standardabweichung von $\sigma = 100$. Wie groß ist dann die Wahrscheinlichkeit $P = P(400 \leq X \leq 500)$ dafür, dass eine Ausprägung des Merkmals den Wert 400 nicht unter- und den Wert 500 nicht überschreitet, also in das Intervall $[400; 500]$ fällt?

Geben Sie das Ergebnis auf *vier* Stellen nach dem Dezimalkomma genau an. Verwenden Sie für das **Dezimalkomma** wieder ein **eigenes Feld**.

(numerisch) $P =$ | | | | | | | |

Aufgabe 47 (lineares Regressionsmodell / KQ-Schätzung) (3 Punkte)

Für zwei Merkmale X und Y liegen Daten $(x_i; y_i)$ aus 6 Beobachtungsperioden vor $(i = 1, 2, ..., 6)$. Aus den Daten wurde $\overline{x} = 14$ und $\overline{y} = 2,5$ errechnet sowie

$$\sum_{i=1}^{6} (x_i - \overline{x})^2 = 40; \qquad \sum_{i=1}^{6} (y_i - \overline{y})^2 = 1,28; \qquad \sum_{i=1}^{6} (x_i - \overline{x})(y_i - \overline{y}) = -6,8.$$

Wenn man für die Beobachtungspaare $(x_i; y_i)$ unterstellt, dass zwischen x_i und y_i ein linearer Zusammenhang besteht, kann man diesen durch das Regressionsmodell $y_i = \alpha + \beta x_i + u_i$ beschreiben und mit den obigen Informationen die Regressionskoeffizienten nach der Methode der kleinsten Quadrate schätzen, also eine Regressionsgerade $\widehat{y} = \widehat{\alpha} + \widehat{\beta} x$ bestimmen.

Berechnen Sie den Wert $\widehat{\alpha}$, also den Schnittpunkt der Regressionsgeraden mit der y-Achse. Tragen Sie Ihr Ergebnis auf *drei* Stellen nach dem Dezimalkomma genau rechtsbündig in das Antwortfeld ein. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch) $\widehat{\alpha} =$ | | | | | | | |

Aufgabe 48 (Bestimmung des kritischen Werts bei einem Test) (3 Punkte)

Es seien $n = 30$ unabhängige Beobachtungen für ein normalverteiltes Merkmal gegeben (Normalverteilung mit unbekanntem Erwartungswert μ). Die Varianz σ^2 bzw. die Standardabweichung σ werde nicht als bekannt vorausgesetzt. Vielmehr liege Information über σ nur in Form einer Schätzung $\widehat{\sigma} = S^*$ vor. Getestet werden soll

$$H_0: \ \mu \leq \mu_0 \qquad \text{gegen} \qquad H_1: \ \mu > \mu_0$$

zum Signifikanzniveau $\alpha = 0,05$. Als Prüfstatistik wird man dann die Prüfgröße

$$T = \frac{\overline{X} - \mu_0}{\widehat{\sigma}} \cdot \sqrt{n} = \frac{\overline{X} - \mu_0}{S^*} \cdot \sqrt{n}$$

für den Test heranziehen und die Nullhypothese verwerfen, wenn der aus den Daten errechnete Wert der Prüfgröße einen bestimmten kritischen Wert überschreitet. Bestimmen Sie diesen kritischen Wert.

Tragen Sie Ihr Ergebnis auf *drei* Stellen nach dem Dezimalkomma genau rechtsbündig in das Antwortfeld ein. Verwenden Sie für das **Dezimalkomma** ein **eigenes Feld**. Vergessen Sie nicht, Ihre Antwort rechtzeitig vor dem Ende der Klausur auf den Markierungsbogen zu übertragen.

(numerisch)
